Storage and Data Movement

Where we are

Industry is beginning to pay attention to our sort of complex storage problem.

Highly mixed local and distant access, almost without predictability or controls, is still a bit alien to vendors.

Meanwhile, experiments have built layers of subscription and transfer management with badly engineered feedback and channel interaction.

What we have

dCache: Fermilab + DESY + other, newer collaborators.

Enstore: Fermilab + PIC + interested parties.

SRM on dCache.

7391 terabytes on tape. This is R/W data.

60+ Gb/s of offsite bandwith.

The hard parts

A storage service must protect itself from offered load.

“Fair share” is rarely the right approach.

“First-come, first-served” fails in quite other ways.

Platform stress

A mixture of high-speed long-distance peers and low-speed local clients in vast numbers is very hard on kernel resources.

Running a highly threaded, bottleneck-sensitive application on such a stressed platform is a challenge.

Approaches

1. Handle a few transfers at a time, controlling every aspect to achieve heroic performance.

A couple of bad peers can ruin your day.

2. Massive parallelism, low performance per transfer with high aggregate.

Hot spots, probable over-provisioning.